

Anomaly detection in clutter using spectrally enhanced Ladar

Puneet S. Chhabra^a, Andrew M. Wallace^a and James R. Hopgood^b

^aSchool of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK

^bSchool of Engineering, University of Edinburgh, Edinburgh, UK

ABSTRACT

Discrete return (DR) Laser Detection and Ranging (Ladar) systems provide a series of echoes that reflect from objects in a scene. These can be first, last or multi-echo returns. In contrast, Full-Waveform (FW)-Ladar systems measure the intensity of light reflected from objects continuously over a period of time. In a camouflaged scenario, e.g., objects hidden behind dense foliage, a FW-Ladar penetrates such foliage and returns a sequence of echoes including buried faint echoes. The aim of this paper is to learn local-patterns of co-occurring echoes characterised by their measured spectra. A deviation from such patterns defines an abnormal event in a forest/tree depth profile. As far as the authors know, neither DR or FW-Ladar, along with several spectral measurements, has not been applied to anomaly detection. This work presents an algorithm that allows detection of spectral and temporal anomalies in FW-Multi Spectral Ladar (FW-MSL) data samples. An anomaly is defined as a full waveform temporal and spectral signature that does not conform to a prior expectation, represented using a learnt subspace (dictionary) and set of coefficients that capture co-occurring local-patterns using an overlapping temporal window. A modified optimization scheme is proposed for subspace learning based on stochastic approximations. The objective function is augmented with a discriminative term that represents the subspace's separability properties and supports anomaly characterisation. The algorithm detects several man-made objects and anomalous spectra hidden in a dense clutter of vegetation and also allows tree species classification.

Keywords: Ladar, LiDAR, ATR, full-waveform, anomaly detection, clutter modelling, feature extraction, sparse representation, multi-spectral, subspace learning, dictionary learning

1. INTRODUCTION

Laser detection and ranging (Ladar) is an active depth (spatial) sensing modality, operated from ground, aerial and space-based platforms. This has found a variety of applications in remote sensing and military scenarios for its ability to extract spatial information from a 3D scene. Passive Multispectral (MSI) and Hyperspectral Imaging (HSI) with dozens and hundreds of wavelengths respectively, measure spectral signatures for each individual image pixel with no spatial information. A fusion of spectral imagery with 3D spatial information may seem desirable. This comes at a cost of both processing time and complexity. The evolution from passive MSI and HSI into spectrally aware active Laser sensors seems natural since simultaneous measurements can eliminate data synchronization errors (both in space and time). Full-waveform Multispectral Ladar (FW-MSL) signals are active spatial measurements that are made simultaneously at different wavelengths using Ladar based sensors. In the past, Ladar sensing has seen applications in ecological studies for tree species classification and segmentation,¹ geo-morphology² and foliage filtering.³ The relevance of urban scene classification with applications to remote sensing has also been studied extensively.³⁻⁵

More recently, small foot-print sensing using MSL has been possible in retrieving structural and physiological properties of vegetation.⁶ Although popular with the remote sensing community, the choice of MSL sensors and raw waveforms for anomaly detection in complex environments (urban city or dense forest) has not been investigated. Echo characterisation and classification is an active area of research but this requires sophisticated echo modelling algorithms which operate on the entire dataset. The aim here is to propose a framework that is smarter in its search for anomalies.

Further author information: (Send correspondence to Puneet S. Chhabra.)
P.C: E-mail: psc31@hw.ac.uk, Telephone: +44 131 451 4168

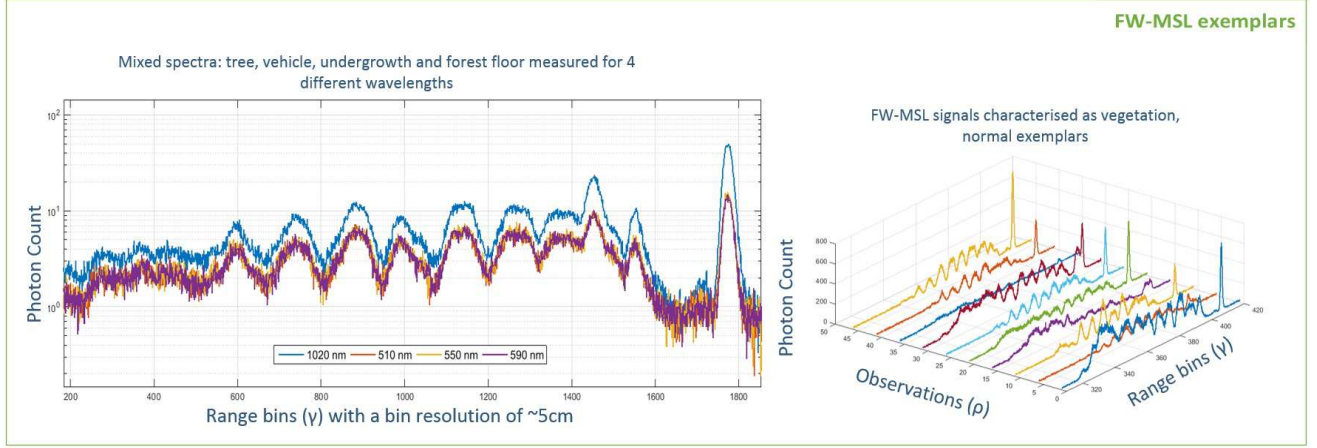


Figure 1. Example FW-MSL signals used to detect anomalies. (Left) A typical FW-MSL return of a large foot-print with a range resolution of $\sim 5\text{cm}$ per bin. The backscattered returns shown here are recorded for 1020, 510, 550 and 590 nm wavelengths. This depth profile is a mixture of different spectra: vehicle, conifer, undergrowth and bare-ground. When observed from an altitude the tree top is the first peak, 600^{th} bin on the left. (Right) Stacked FW-MSL exemplars.

Automatic Target Recognition (ATR) are scenarios that may benefit, for example, in automatic detection of known targets hidden or camouflaged under dense foliage. The nature of anomaly detection differs from the ATR notion. The type of target, its structure and material may be unknown. Consider a surveying scenario where a forest is scanned looking for normality and abnormality in what is being perceived through different sensor modalities. A set of measurements can be made on what is being observed and one may be interested in narrowing down the target search window by finding spatial and spectral anomalies for further structural analysis. This can be carried out in two phases:

- *Phase 1* of the survey operates on a *large-footprint* setting which aims to identify regions with spectral and spatial (time) anomalies, i.e., similar spectral signatures found at depths which do not comply with learnt models. This paper proposes a method to support Phase 1. Once these regions are identified, Phase 2 is applied to regions that contain such anomalies.
- *Phase 2* A *small-footprint* scanning scheme can be employed along with expensive signal modelling algorithms to deliver a dense 3D point cloud for structural analysis.

In order to support Phase 1 and Phase 2, Ladar systems have been designed and developed operating on simultaneously upto 4 wavelengths.⁷ In order to support future instrument design and parameter selection experiments on simulated and synthetic data have been carried out on a range of wavelengths: 4, 8, 16 and 32. Figure 1 illustrates example FW-Ladar signals measured over a plot of vegetation.

1.1 Contributions

In this paper, a two-stage surveillance framework involving a sensing scheme that makes use of both small and large footprint Ladar scanning is proposed. In order to model natural trees or forest plots, a layered representation is considered that conforms to a sparse approximation and discrimination model. The framework includes:

- This work aims to avoid computationally expensive signal modelling algorithms in order to characterize backscattering echoes for target localisation on large FW-MSL datasets. To support such an aim, a cueing algorithm that reduces target search space by identifying anomalies in a large urban or forest scene is proposed. Such an approach is novel since it can be an efficient way of combining large and small footprint scanning, especially when surveying large environments and the multi-spectral nature of the sensors used.

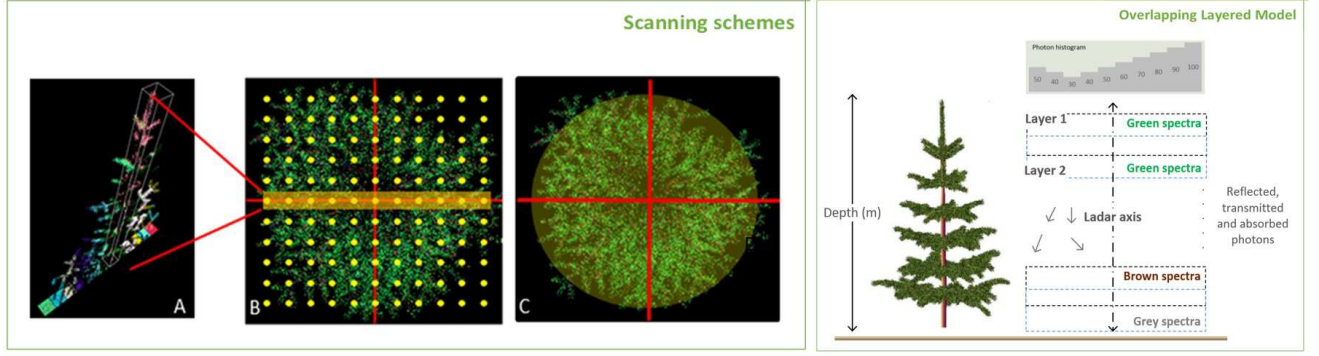


Figure 2. (A) A cross-section of a single conifer tree. (B) A *push-broom* style (small-foot print) dense scan pattern resulting in a cross-section slice of a tree. A *flash* style (large-foot print) scan pattern is illustrated in (C). (D) The forest/tree layered model.

- A semi-supervised subspace (a set of linearly dependent basis, i.e., dictionary) learning method (Section 3) which learns normal Ladar signals of different conifer species and forest terrain comprising of 8 different spectra (bark, needle, terrain, under-growth, steel, concrete, brick and clay tiles); and each new observation is transformed and reconstructed using the learnt dictionary. High reconstruction errors reveal hidden anomalies. The optimisation routine used here is influenced from the compressed sensing literature except that objective function is modified in order to classify tree species and other signals. See Algorithm 2.
- The proposed framework can emphasize any unknown as opposed to a known signature⁸ both in spatial and spectral context. To the best of authors knowledge, this paper is the first attempt at applying such techniques to foliage penetration and anomaly detection on raw FW-MSL measurements.
- This work also shows that the proposed approach is widely applicable as the definition of an anomaly is not restrictive to a particular target type. The concept of context is associated with different FW-MSL measurements across different layers. In the past, this has been over-looked, echo modelling schemes have treated each individual echo independently and characterise with its location and amplitude. This work shows that relationship learning across layers and wavelengths can discover co-occurring patterns which can be highly relevant to anomaly detection.

1.2 Outline

An outline of the paper is as follows: Section 2 provides a mathematical formulation of the problem and description of an anomaly. Section 3 offers a solution to the problem with experiments and results on the data in Section 4. Finally, Section 5 provides conclusions and future work.

2. PROBLEM DEFINITION

Anomaly detection can be defined as a problem of detecting data samples that significantly deviate from an expected model. Methods that can capture essential patterns both in temporal and spectral dimensions within the original data space are desirable. Such patterns contribute significantly to the generating cause of an observation, a measured signal, and such representation are in a lower dimensional space.

Recent developments in *signal approximation*⁹ and dimensionality reduction have an underlying concept, *Sparsity*, which means that an original signal, which is dense in certain basis, once its is transformed in to another convenient basis i.e., the contribution vector is almost sparse. In this paper, it is assumed that an expected prior conforms to a sparse representation model, where the coefficients represent the contribution of a set of *atoms*, together called a subspace. This work also employs method that extract and learns such a subspace and their coefficients from raw FW-MSL data. Together, they contribute towards an *accurate* representation of the original FW-MSL measurement data. The following section introduces a layered representation of a forest, followed by an FW-MSL approximation model.

2.1 Layered representation of forest/tree

An explicit assumption is made here that a depth profile through a dense forest or an individual tree can be represented by a series of instrumental Ladar returns from a set of *layers* (see Figure 2D) at different ranges. Thus, layer position defines the *context* of a tree and the number of layers, the number of individual spectra measured within a layer, and their contribution defines its *behaviour*.

A depth profile of a tree or forest are made up of countless surface responses from leaf, bark and other man-made surfaces found at varying depths. This is a *leaf-on* condition, dense healthy forests, mostly during summer. During winter, a *leaf-off* condition, most of the laser light penetrates trees quiet easily. Such a condition is overlooked at this stage. The illustration in Figure 2 highlights one such example with spectral response registered at each layer. The top-most layer records green, since it mostly consists of high chlorophyll content representing natural material. The bottom-most layer records brown dead leaves and undergrowth and finally forest floor records spectra different to other layers, color coded here as grey. This sequence clearly deviates under diverse vegetation or with varying tree heights. Hence, the intuition here is to capture the relationship between overlapping layers in order to capture any co-occurring patterns.

This can be achieved by breaking down a data sample, in its entirety in to several neighbouring sub-sequences and measure the co-occurrence of local-patterns within every sub-sequence. For the sake of simplicity, running an overlapping sliding window over wavelength(λ)-range(γ) series resulting in multiple sub-sequences. The aim is to learn lower-dimensional subspaces in order to seek unusual local patterns that do not appear in reference range-amplitude series. An implicit assumption made here is that the flying altitude is known and that each individual FW-MSL measurement is registered with respect to the flying altitude, i.e., that the length of the two samples and the range bin index are same. Next, the concept of normality and abnormality in the context of range-amplitude-wavelength, FW-MSL data, is explained.

2.2 The concept of normality and abnormality

In a semi-supervised anomaly detection technique, it is assumed that a database of *normal* exemplars that conform to non-anomalous data is known. Within the realm of this work, these exemplars may belong to several different classes, i.e., different tree species or Ladar returns of vegetation and bare forest floor. A continuous range-amplitude depth profile of a tree or forest is made up of several different spectra: bark, needle, dead leaves, undergrowth and forest floor. Such data samples are qualified as *normal*, free from any man-made objects. Deviation from such learnt model qualifies as *abnormality*.

Scenes observed using active/passive range sensors are made up of objects of interests, background clutter, sensor and environmental noise. Background clutter can be defined as something that is neither anomalous nor an object of interest. Interestingly, an anomaly characterised by a context or an outcome of a condition can be referred as a *contextual anomaly*, or a *conditional anomaly*. In the context of FW-Ladar measurements, the input data is a range-amplitude-wavelength series and the context here is occurrence of an event at a particular range γ or wavelength λ .

In a layered representation (Figure 2) an anomaly can occur in any range bin, i.e., depth. This could be due to the variability in tree height across a section of the forest, tree health (climatic conditions) or a more interesting cause: presence of man-made object(s). It is highly impractical to collect data samples for all abnormal cases, but, what is available in abundance is large amount of training data, normal events only, plots of natural forests with trees. The anomaly detection problem can be defined as follows:

One wishes to learn the abundance of distinctive material spectra and their relationships across different layers within a range-amplitude series. The aim is to detect a deviation from the learnt normal models resulting in an abnormal event.

2.3 Waveform Model

A set of Ladar signals measured for an individual wavelength λ can be defined as:

$$L_{\lambda} = \{l_{\lambda,1}, ..., l_{\lambda,\Gamma}\}, \quad (1)$$

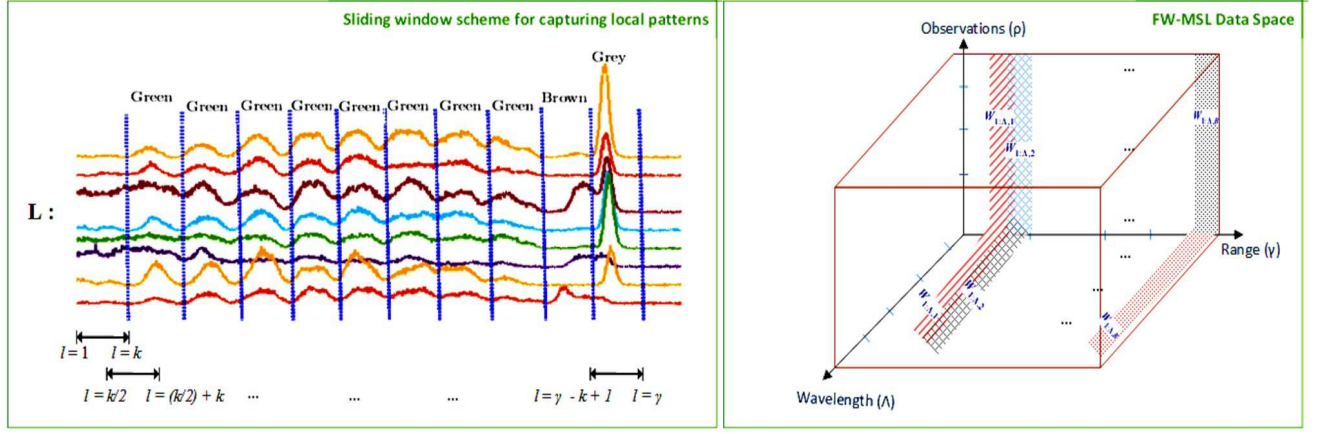


Figure 3. Sliding window concept in order to extract over-lapping range-amplitude series.

where Γ is the number of of range bins, i.e., $\gamma \in [1, 2, \dots, \Gamma]$ and $\lambda \in [1, 2, \dots, \Lambda]$. Formally, $l_{\lambda, \gamma}$ expresses the photon count within the γ^{th} channel (range bin) of the λ^{th} spectral band being observed. Here, the assumption is that each photon count $l_{\lambda, \gamma}$ is drawn independently from a Poisson distribution:

$$Prob(L_{\lambda}) = e^{-\delta_{\lambda, \gamma}} \frac{(\delta_{\lambda, \gamma})^{l_{\lambda, \gamma}}}{l_{\lambda, \gamma}!}, \quad (2)$$

In 2, $\delta_{\lambda, \gamma}$ is defined as:

$$\delta_{\lambda, \gamma} = s_{\lambda, \gamma} K_{0, \lambda}(\gamma) + b_{\lambda}, \quad (3)$$

where $s_{\lambda, \gamma}$ denotes different surfaces a reference beam interacts with, $K_{0, \lambda}(\cdot)$ is the photon impulse response, and b_{λ} is background and dark photon level across all bins. The aim of this work is not to model or deconvolve such responses in order to find echoes. Instead, the proposed approach aims to operate on raw FW-MSL signals collected over a synthetic forest in order to leave sophisticated echo modelling algorithms to operate on only a sub-set of *interesting* regions in the scene.

2.4 Mapping raw Ladar data into sub-spaces

FW-MSL signals measured for Λ wavelengths can be expressed as:

$$\mathbf{L} = [L_1, \dots, L_{\Lambda}]^T. \quad (4)$$

The original range-amplitude-wavelength series \mathbf{L} for different wavelengths (Λ) is mapped on to a set of over-lapping sub-sequences \mathbf{W} and then transformed in to a feature subspace \mathbf{F} . The mapping can be written as follows:

$$\mathbf{L} \rightarrow \mathbf{W} \rightarrow \mathbf{F} \quad (5)$$

The first half of the mapping is explained below followed by the final mapping into \mathbf{F} is explained in more detail at the end of Section 3.1 using 14 and 15.

Overlapping sub-sequence representation ($\mathbf{L} \rightarrow \mathbf{W}$) Figure 3 illustrates the windowing concept and subspace representation of normal vs abnormal events occurring in the range-amplitude-wavelengths space. $\mathbf{L} \in \mathbb{R}^{\Gamma \times \rho \times \Lambda}$ for ρ training samples, Λ number of wavelengths and γ number of range bins. So the $\mathbf{L} \rightarrow \mathbf{W}$ can be written as:

$$\mathbf{W} = [W_{\lambda}, \dots, W_{\Lambda}]^T, \quad (6)$$

where W_{λ} is a sub-sequence representing a layer of the forest/tree of size $B = \frac{k}{\Gamma}$ that are partitioned on the range (γ -axis), where k controls the window size and m is the amount of overlap. For a fixed wavelength λ this

can be written as:

$$W_\lambda = \left\{ \underbrace{l_{\lambda,1}, \dots, l_{\lambda,k}}_{W_{\lambda,1}}, \underbrace{l_{\lambda,1+m}, \dots, l_{\lambda,k+m}}_{W_{\lambda,2}}, \dots, \dots, \underbrace{l_{\lambda,\Gamma-k+1}, \dots, l_{\lambda,\Gamma}}_{W_{\lambda,B}} \right\} \quad (7)$$

Subspace representation Each data sub-sequence, $W_{\lambda,i}$, where $i = [1, 2, \dots, B]$ is the sequence index in 7 can be described with an over-complete basis (dictionary or subspace), $D_{\lambda,i}$, a coefficient matrix $C_{\lambda,i}$ plus additive constant background noise:

$$W_{\lambda,i} \approx D_{\lambda,i} C_{\lambda,i}, \quad (8)$$

where $W_{\lambda,i} \in \mathbb{R}^{k \times \rho}$ and $D_{\lambda,i} \in \mathbb{R}^{k \times \eta}$ is an over-complete ($\eta \gg k$) matrix and the resulting coefficient matrix $C_{\lambda,i} \in \mathbb{R}^{k \times \eta}$. The goal here is to find an optimal subset, called *Subspace (Dictionary)* $(D_{\lambda,i})'$ such that the set $W_{\lambda,i}$ can be accurately reconstructed by $(D_{\lambda,i})'$ and the size of this subspace is as small as possible. A constant background Poisson noise model in $W_{\lambda,i}$ is assumed. A simple way of doing this is to randomly and uniformly select basis vectors to build the subspace. This strategy can be risky as it may not capture the variability in the training data. A principled way of selection which selects an optimal subset of $D_{\lambda,i}$ as a subspace is presented below.

Using all the vectors of $D_{\lambda,i}$, denoted by $\{d_j\}_{j=1}^\eta$ (see Section 3.1), all vectors of $C_{\lambda,i}$, denoted by $\{C_{\lambda,i}\}_{j=1}^\eta$ a signal $W_{\lambda,i}$ can be represented as a sparse linear combination of these columns, referred to as *atoms*. The representation of $W_{\lambda,i}$ may either be exact (8) or an approximation $\|W_{\lambda,i} - D_{\lambda,i} C_{\lambda,i}\|_2 \leq \epsilon$. For a full-rank matrix $D_{\lambda,i}$ and $k < \eta$, there are an infinite number of solutions to this representation (8). A stop constraint is needed on this solution, either set a sparsest approximation constraint, i.e. $\min_{C_{\lambda,i}} \|C_{\lambda,i}\|_0$, where $\|\cdot\|_0$ is the l^0 quasi-norm, which counts the non-zero entries of a vector or stop when the error term ϵ stops decreasing and goes below a threshold. An optimal can be defined by minimising the following objective function:

$$\arg \min_{D_{\lambda,i}, C_{\lambda,i}} [\|W_{\lambda,i} - D_{\lambda,i} C_{\lambda,i}\|_2^2 + \beta_1 \|C_{\lambda,i}\|_0], \quad (9)$$

where the parameter $\beta_1 > 0$ is a regularization scalar that balances the trade-off between the reconstruction accuracy and sparsity. Finding a solution to (9) is an NP-hard problem primarily due to its combinatorial optimisation, i.e., minimising for best reconstruction with least number of basis vectors are convex individually but not convex jointly. This can be solved iteratively.

A suboptimal solution to this problem is found by replacing the l_0 norm in (9) with the l_1 norm and re-write (9) using components of $C_{\lambda,i}$ as:

$$\arg \min_{D_{\lambda,i}, C_{\lambda,i}} \left[\|W_{\lambda,i} - D_{\lambda,i} C_{\lambda,i}\|_2^2 + \beta_1 \sum_{j=1}^\eta \|\{C_{\lambda,i}\}_j\|_1 \right], \quad (10)$$

where $\|\cdot\|_1$ is the l_1 norm. In the following section an iterative way to solve 10 is explained in more detail.

3. ADOPTED APPROACH

A solution to (10) leads to a good approximation of measurement matrix $W_{\lambda,i}$. A discriminating trait is added to the original problem and finally the reconstruction error is used to differentiate normal from abnormal data. This framework also allows for classifying of test ladar signals in to different classes (e.g., tree species). A natural forest environment is made up of several tree species with variable tree height. Hence, one cannot fix one tree *layer* model for different tree species. One solution to capture such variability is to have species bound subspaces and coefficients.

As illustrated in the previous section the data is layered in to overlapping sub-sequences. The training step is shown in the form a pseudo code in Algorithm 1. The optimisation routine with an discriminative objective function is presented in Algorithm 2. For the sake of completeness the optimisation routine is presented in Algorithm 2. The complete approach can be briefly listed in three steps:

Algorithm 1 Anomaly Training

Input: Configuration file with $\Gamma, \rho, \Lambda, k, m, weights, mode$ and η

Output: Learnt dictionaries and coefficients, i.e., $\mathbf{L} \rightarrow \mathbf{W} \rightarrow \mathbf{F}$

```
1: procedure ANOMALYTRAIN(configFile)
2:   normalise  $\mathbf{L}$ , FW-MSL data
3:   for  $i = 1$  to num_of_Layers do
4:     Map  $\mathbf{L} \rightarrow \mathbf{W}$  using 7
5:     Call:  $[\mathbf{C}_{\lambda,i}, \text{optimized } \mathbf{D}_{\lambda,i}] = \text{SAD}(\mathbf{W}_{\lambda,i}, \mathbf{D}_{\lambda,i}, \text{mode}, \text{weights})$  - Algorithm 2     $\triangleright$  see Section 3.1
6:     Learn local-patterns on joint-wavelength coefficient matrices     $\triangleright$  See section 3.2
7:     Transform layer wise coefficients of  $\mathbf{W}_{\lambda,i}$  to a lower dimensional space     $\triangleright$  See section 3.2
8:     Map  $\mathbf{W} \rightarrow \mathbf{F}$      $\triangleright$  Equation 14 and 15
9:     Save  $\mathbf{F}_{train}$ 
10:   end for
11: end procedure
```

1. *Subspace selecting & Learning* In training mode a matching pursuit method is employed in order to learn optimal subspaces, $(\mathbf{D}_{\lambda,i})'$ and multi-wavelengths coefficient matrix \mathbf{F} , i.e., $\mathbf{W} \rightarrow \mathbf{F}_{train}$. Such a representation can be extended by including a class discriminative (non-anomalous) criterion in to the objective function such that during the approximation stage discriminating subspaces can be learned for different normal classes (conifer species, bare forest floor, and undergrowth). In testing mode, using the learnt subspace a new coefficient matrix \mathbf{F}_{test} is computed.
2. *Relationship learning & anomaly detection* \mathbf{F}_{train} is decomposed using a semantic learning technique, transformed and reconstructed. A test sample is projected back to the subspace and reconstructed, a high reconstruction error reveals the anomalies present in the test data samples.

3.1 Step 1: Learning discriminating subspaces and Coefficients

When solving (10) The aim here is to approximate a subspace that allows an assumption is made that the training data are labelled exemplars for only the normal class. Equation (10) can be updated with a discrimination function parametrised by $D_{\lambda,i}$ as a Fisher Discriminant which has been very popular with Linear Discriminant Analysis (LDA), which is a generalisation of Fisher's linear discriminant.¹⁰ The discrimination power is maximised when the spatial distribution of samples from similar classes (scatter matrix, S_w) are closer to each other when compared to samples from different classes (scatter matrix, S_b). For a *subspace*, $D_{\lambda,i} = [d_1, d_2, \dots, d_\eta]$ (where d_1, d_2, \dots, d_η are the *atoms* in the subspace, 6), of which η_i samples are in class Ω_i , for $1 \leq i \leq \Omega$, mean is defined as μ_i and variance v_i^2 for class Ω_i as $\mu_i = \frac{1}{\eta_i} \sum_{d \in \Omega_i} \mathbf{d}$, and $v_i^2 = \frac{1}{\eta_i} \sum_{d \in \Omega_i} \|\mathbf{d} - \mu_i\|_2^2$. The mean of all samples can be written as: $\mu = \frac{1}{\eta} \sum_{i=1}^{\eta} d_i$. Finally, the Fisher discrimination function for the proposed objective function is defined as:

$$K(D_{\lambda,i}) = S_w^{-1} S_b, \quad (11)$$

where S_w , the *inner-class* scatter matrix and S_b , the *intra-class* scatter matrix are defined as follows:

$$S_b = \left\| \sum_{i=1}^{\Omega} \eta_i (\mu_i - \mu) (\mu_i - \mu)^T \right\|_2^2, \quad S_w = \sum_{i=1}^{\Omega} v_i^2 \quad (12)$$

Using (12) update (10) with:

$$\arg \min_{D_{\lambda,i}, C_{\lambda,i}} \left[K(D_{\lambda,i}) + \beta_1 \sum_{j=1}^{\eta} \|\{C_{\lambda,i}\}_j\|_1 + \beta_2 \sum_{j=1}^{\eta} \|\mathbf{W}_{\lambda,i} - D_{\lambda,i} \{C_{\lambda,i}\}_j\|_2^2 \right], \quad (13)$$

where β_1 and β_2 are positive scalar chosen as a trade-off between reconstruction error, coefficients contribution and discrimination. An iterative optimisation routine motivated by pursuit algorithms^{11,12} is suggested in order

Algorithm 2 Simultaneous Approximation & Discrimination (SAD)

Input: Measurement vector $W_{\lambda,i} \in \mathbb{R}^{k \times \rho}$; number of atoms, η ; *weights* = $[\beta_1, \beta_2]$, and mode

Output: Learnt dictionary $(D_{\lambda,i})'$; Coefficients $C_{\lambda,i}$

```

1: procedure SAD( $(\mathbf{W}_{\lambda,i}, \mathbf{D}_{\lambda,i}, \text{mode}, \text{weights})$ )
2:   Initialize: Index set  $\omega = \phi$  and residual  $(R_{\lambda,i})_0 = W_{\lambda,i}$ 
3:   while  $R_{\lambda,i} \neq 0$  do
4:     Identify: An atom from subspace  $D_{\lambda,i}$  that maximizes:
5:     if  $\text{mode} == \text{approximate\_only}$  then
6:        $d : \arg \max_{d \in D_{\lambda,i}} J(d^T(R_{\lambda,i})_\eta, \beta_1)$  ▷ Equation 10
7:     else
8:        $d : \arg \max_{d \in D_{\lambda,i}} J(d^T(R_{\lambda,i})_\eta, \beta_1, \beta_2)$  ▷ Equation 13
9:     end if
10:    Compute: Orthogonal projection  $(P_{\lambda,i})_\eta$  (for remaining atoms):
11:     $(P_{\lambda,i}) \in \{d\}_{j=1}^{\eta-1} = D_{\lambda,i} * \text{inv}((D_{\lambda,i})^T * D_{\lambda,i}) * (D_{\lambda,i})^T$ 
12:    Compute: The new residue is given as:
13:     $(R_{\lambda,i})_\eta = W_{\lambda,i} - (P_{\lambda,i})_\eta W_{\lambda,i}$ 
14:    Increment:  $\eta = \eta + 1$ 
15:  end while
16: end procedure

```

to solve 10. It is coined as *Simultaneous Approximation & Discrimination (SAD)* in this work as a discriminating term is added to the original objective function.*

The algorithm in its entirety is illustrated above. Repeat step 2 for each individual layer (sub-sequence) and stack the coefficient matrix $C_{\lambda,i}$, where $\lambda = [1, 2, \dots, \Lambda]$ (output of Algorithm 2). So, for layer 1 $F_{\Lambda,1}$ matrix can be expressed as:

$$F_{\Lambda,1} = \begin{bmatrix} C_{\lambda,1} \\ \vdots \\ \vdots \\ C_{\Lambda,1} \end{bmatrix} \in \mathbb{R}^{(\Lambda \times \rho) \times \eta}, \quad (14)$$

and a set of such matrices can be represented as:

$$\mathbf{F}_{train} = \left\{ \underbrace{F_{\lambda,1}}_{W_{\lambda,1}}, \underbrace{F_{\lambda,2}}_{W_{\lambda,2}}, \dots, \dots, \underbrace{F_{\Lambda,B}}_{W_{\Lambda,B}} \right\}. \quad (15)$$

The label under-set each $F_{\Lambda,B}$ shows which layer it belongs to. Conceptual diagram showing how \mathbf{F}_{train} matrix is constructed is shown in Figure 4. The final step is to reduce the dimensionality in the Λ dimension and learn the relationships across different layers at the same time. This is shown below.

3.2 Step 3: Relationship Learning & Anomaly Detection

The Singular Value Decomposition (SVD) is used to decompose F_{train} :

$$\mathbf{F}_{train} = U \Sigma V^T, \quad (16)$$

where, $U = [\mathbf{u}_1, \dots, \mathbf{u}_{\Gamma \times \rho}]$ (ρ = number of training examples) and $V = [\mathbf{v}_1, \dots, \mathbf{v}_\eta]$ (η = number of basis vectors) are sets of singular vectors, and Σ , a diagonal matrix of the singular values. In this representation, the vectors of U correspond to the elementary atoms that make up a radar signal and vectors of V corresponds to a

*If the classification mode is selected replace (10) by (13)

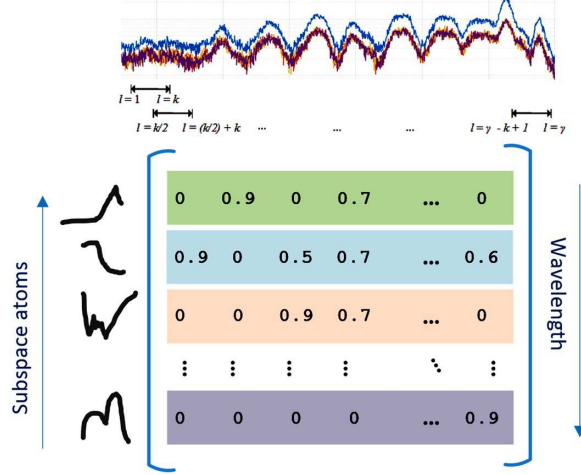


Figure 4. Conceptual diagram of multi-wavelength *atom-layer* relationship matrix (coefficient matrix) which is learnt for several over-lapping sub-sequences (columns) on range-amplitude ladar signals.

tree/forest layer (window) in the semantic space. To preserve only q essential semantic relationships, one needs to approximate the *atom-layer* relationship matrix as follows:

$$\mathbf{F}_{train} \approx U_q \Sigma_q V_q^T. \quad (17)$$

In the training phase apply (17) to the atom-layer (coefficient) matrix \mathbf{F}_{train} resulting in U_q and Σ_q . During testing, the atom-layer coefficient matrix \mathbf{F}_{test} is transformed in to a lower-dimensional semantic space using:

$$\hat{\mathbf{F}}_{test} = \Sigma_q^{-1} U_q^T \mathbf{F}_{train}, \quad (18)$$

and reconstructed in to the original space using:

$$\tilde{\mathbf{F}} = U_q \Sigma_q \hat{\mathbf{F}}_{test}. \quad (19)$$

This step completes the mapping of $\mathbf{W} \rightarrow \mathbf{F}$ which results in a mapping $\mathbf{F}_{train} \in q \times \eta$ where, $q \ll (\Lambda \times \rho)$.

Anomaly Detection and Score It should be noted that while the low-rank nature of U_q and Σ_q captures only the essential semantic relationships, (19) cannot reconstruct the original atom-layer coefficient matrix perfectly. The underlying intuition here is that if the co-occurrence relationships in the test ladar samples are different to those in the training data set, the reconstruction error will be large, i.e., the anomalous elements of the coefficient matrix will have significant large value. This is computed by calculating the distance between the learnt representation and reconstructed signal. The reconstruction error on each individual layer is computed:

$$\hat{A}_{score} = \|\mathbf{F}_{train} - \tilde{\mathbf{F}}\|_2^2, \quad (20)$$

Equation (20) computes a score with regard to an individual layer (sliding window), an anomaly score for each FW-MSL signal with N layers can be written as:

$$A_{score} = \frac{1}{N} \sum_1^N \tilde{A}_{score} \quad (21)$$

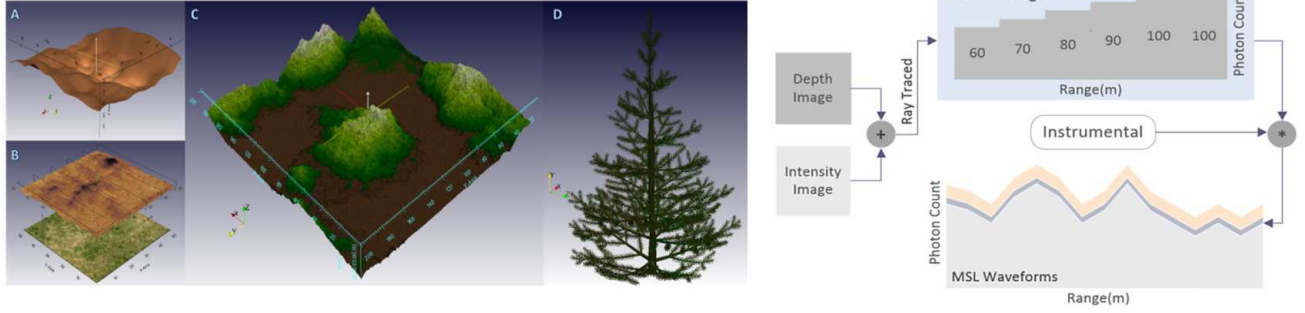


Figure 5. (Left) Natural terrain generation method using fractal based Brownian Motion Fields. In (A) A smooth desert terrain with a high Hurst index H . The two-step terrain generator process is shown in (B), layer 1 is a factorial Brownian field (fBF) and layer 0 is a texture layer that is mapped to the elevation model. (C) A combination of several two-step processes (shown in (B)) in creating a combinatorial terrain with elevated hills and rough terrain. (D) Output of a closed-source tree modeller, Onyxtree that is used in this study for several conifer species. (Right) Once the scene is rendered, it shows how the depth and intensity information can be used to generate synthetic MSL waveforms.

Note If the classification mode is selected, the above steps are followed except that training labels for different classes are made available. In such a scenario, the discriminative objective function (13) is used to learn atom-layer relationships and assign them class labels. For each test sample a K-NN match is computed in order to predict its class label and then compute (18 and 19). Initial experiments show that this reduces false alarms but at an expense of CPU time in solving (13). Due to space constraints, this paper only highlights only anomaly detection scores.

4. SYNTHETIC DATA AND EXPERIMENTS

When scanning an urban setting or dense forest we employ two scanning schemes (See Figure 2); firstly, a *push-broom* style scan where we carry out a dense grid style scan and secondly, a *flash* style scan with larger beam resulting in a large foot-print scan that illuminates an entire tree in one go. In Figure 5 we illustrate the waveform generation process based on depth image and intensity image. A synthetic scene is ray-traced using a custom modified Povray[†] that also generates a depth image. Equation (22) is used to convolve the scene with a Gaussian instrumental (3-5ns) to generate synthetic FW-MSL waveforms. For the sake of simplicity the same instrumental is used across all wavelengths. The received power $P_a(\gamma)$ of a Ladar signal after interacting with S surfaces, at a range R , with D_a as the aperture diameter of the receiver sensor can be expressed as:

$$P_a(\gamma) = \sum_{i=1}^S \frac{D_a^2}{4\pi\lambda^2 R_i^4} P_d(\gamma) * \Gamma(\gamma) * E(\gamma) * \sigma_i(\gamma), \quad (22)$$

where $*$ denotes a convolution operation over $P_d(\gamma)$, the transmitted power, $\Gamma(\gamma)$, the system impulse, $E(\gamma)$, environmental aerosol contribution and $\sigma_i(\gamma)$, is the backscattering cross-section.

4.1 Synthetic Data - Multi-Beam Modeller

The temporal shape of the received signal often referred to as a waveform can be written as a sum of all the echoes from S surfaces. Some of the resulting waveforms for a single-wavelength are shown in Figure 1 with range as its x-axis and amplitude (photon count) on the y-axis. Similar waveforms are generated for 4, 8, 16 and 32 wavelengths.

In this section we first present a method to model 3D forest-terrain and then generate realistic trees using a closed-source popular tree modeller, OnyxTree[‡]. Two sets of datasets are collected, one which contains up to 3 different conifer tree species (Picea Orientalis, Pseudo Suga M2 and Picea Engelmannii) with variable height

[†]<http://www.pov-ray.org>

[‡]<http://www.onyxtree.com>

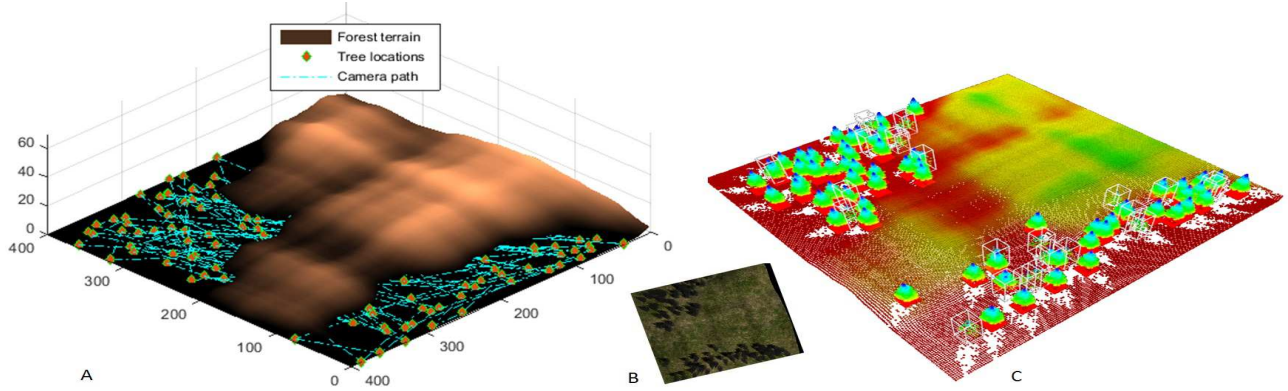


Figure 6. Rendered scene with conifer species. A) Terrain generated along with tree locations and camera path. B) Rendered scene (2D Image). C) 3D point cloud with elevation shaded (light to dark).

and Leaf Area Index (LAI); second, we populate the same environment with man-made objects such as vehicles at random locations. In order to make the setting more realistic we add random undergrowth in between the trees. The reflectance spectra of all the materials used in the experiments (*man-made*[§]: brick, granite, steel, roof-tiles and *natural objects*[¶]: conifer bark, needle and Aridisol (soil)) is illustrated in Figure 7. Next, a brief note on how natural forest terrains are generated is presented.

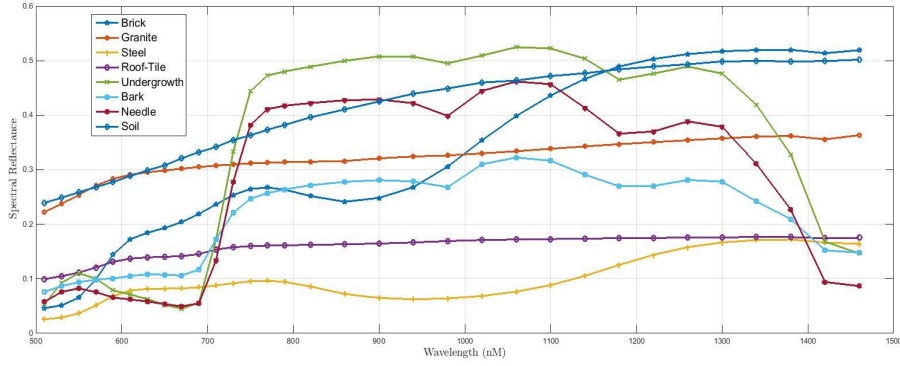


Figure 7. Measured material spectral reflectance for several *natural*: tree bark, needle and soil samples along with data from the Aster spectral library for selected *man-made objects*: building material and galvanized steel.

Natural Terrain Modeller In nature we find natural objects tend to follow a symmetry pattern and self-similarity. Mandelbrot¹⁴ in his influential paper linked such symmetry to a mathematical phenomenon called fractals. Following this method, natural terrains can be modelled using a fractal interface. In this work a forest terrain as a simple continuous-time stochastic process and model it as a fractal or factorial Brownian Field (fBF). The terrains are made up of two layers. Layer one, is an elevation field and texture is mapped to the elevation model in layer two. Figure 5 shows examples of different terrains generated. An elevation function is defined using a generalised Brownian motion without independent increments:

$$E[B_{t_1}, B_{t_2}] = \langle B_{t_1} B_{t_2} \rangle = \frac{\sigma^2}{2} [t^{2H} - (t_2 - t_1)^{2H} + s^{2H}] \quad (23)$$

where B_t is a continuous-time Gaussian Process on $[0, T]$ and σ , The roughness parameter H , called Hurst index is varied in order to control the raggedness of the motion and this can result in a variety of surfaces such

[§]The Aster2.0¹³ spectral library provided by JPL, NASA is used for reflectance values of man-made objects.

[¶]Passive measurements of tree samples under laboratory conditions⁷ was carried out using a stabilized light source and spectrometer, with several needles and bark samples.

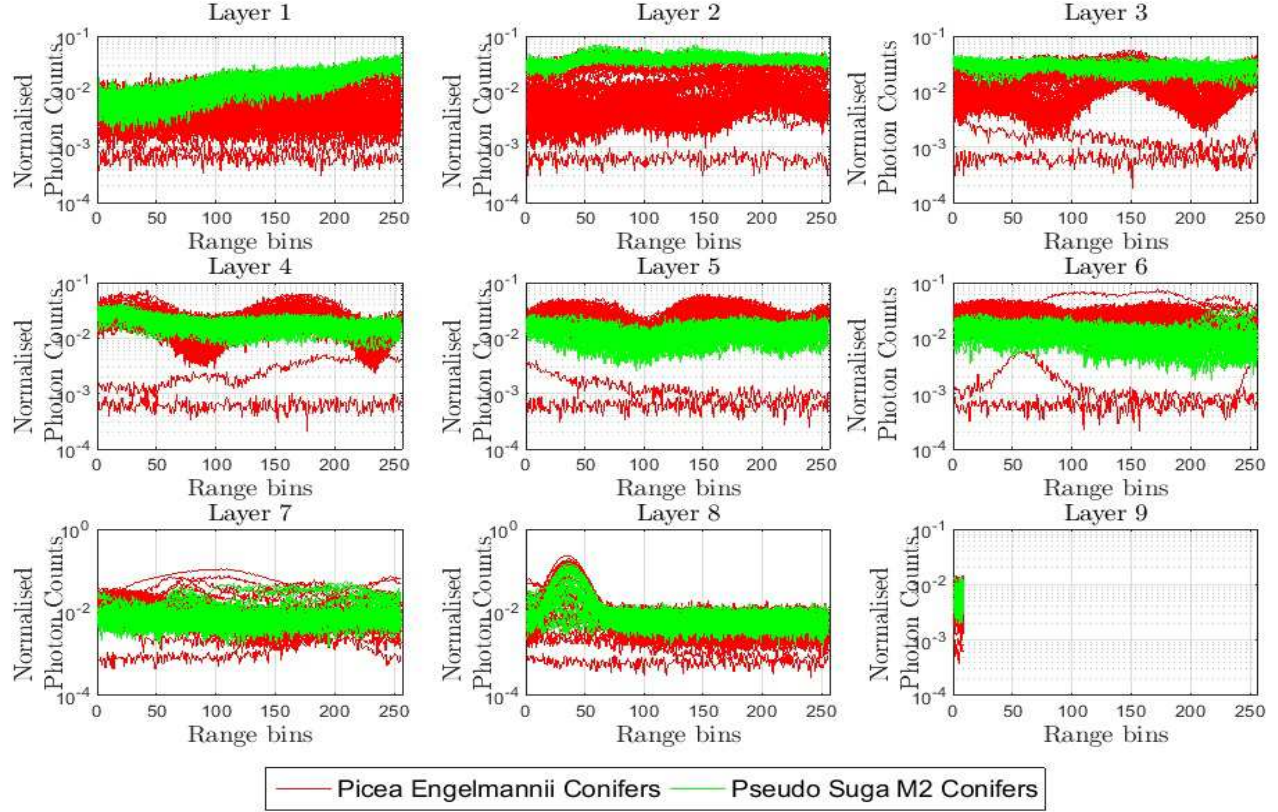


Figure 8. Test FW ladar signals belonging to 2 different conifer species for a single wavelength partitioned in to 9 layers using $\mathbf{L} \rightarrow \mathbf{W}$. In one of the experiments 105 sample belonging to two different conifer species are used. Anomalies are introduced in the form of man-made objects at different depths and layers.

as smooth desert type terrains and mountainous regions. Combinations of more than one surfaces are used to build rocky terrains and also by controlling the shape and geometry of the motion by defining a random moment matrix and an asymmetry property to the matrix. The steps involved are illustrated in Figure 3.2. The terrain grid along with elevation values are generated in MATLAB. Different tree species, man-made objects undergrowth along with generated terrain are then rendered using modified Pov-ray.

Man-made Objects Most of the man-made objects come from manufacturing processes or architectural constructions and are a complex arrangement of basic ensembles (cylinders, spheres or planes). Such objects are used in the experiments in order to study the spectral responses of different surfaces (elevated areas, small objects and vegetation surfaces) with multi-beam and multi-spectral Ladar sensor. Man-made structures with elevated roofs and planar platforms along with vehicles are rendered and placed in a forest setting.

4.2 Results on Synthetic Data

Scenes rendered using the proposed modeller, as illustrated in Figure 6 are used for experimentation. This small plot of forest is made up of two conifer tree species and several man-made objects are placed in between. These objects are vehicles closer to the ground and camouflaged with a random undergrowth introduced in between the trees.

An overhead surveillance flight simulation is assumed. Flight path is traced and several regions of the forest are illuminated using large foot-print waveforms and backscattered returns are recorded using 22. Out of 105 test samples 23 FW-MSL signals contain several anomalies. The detections are shown in Figure 9.

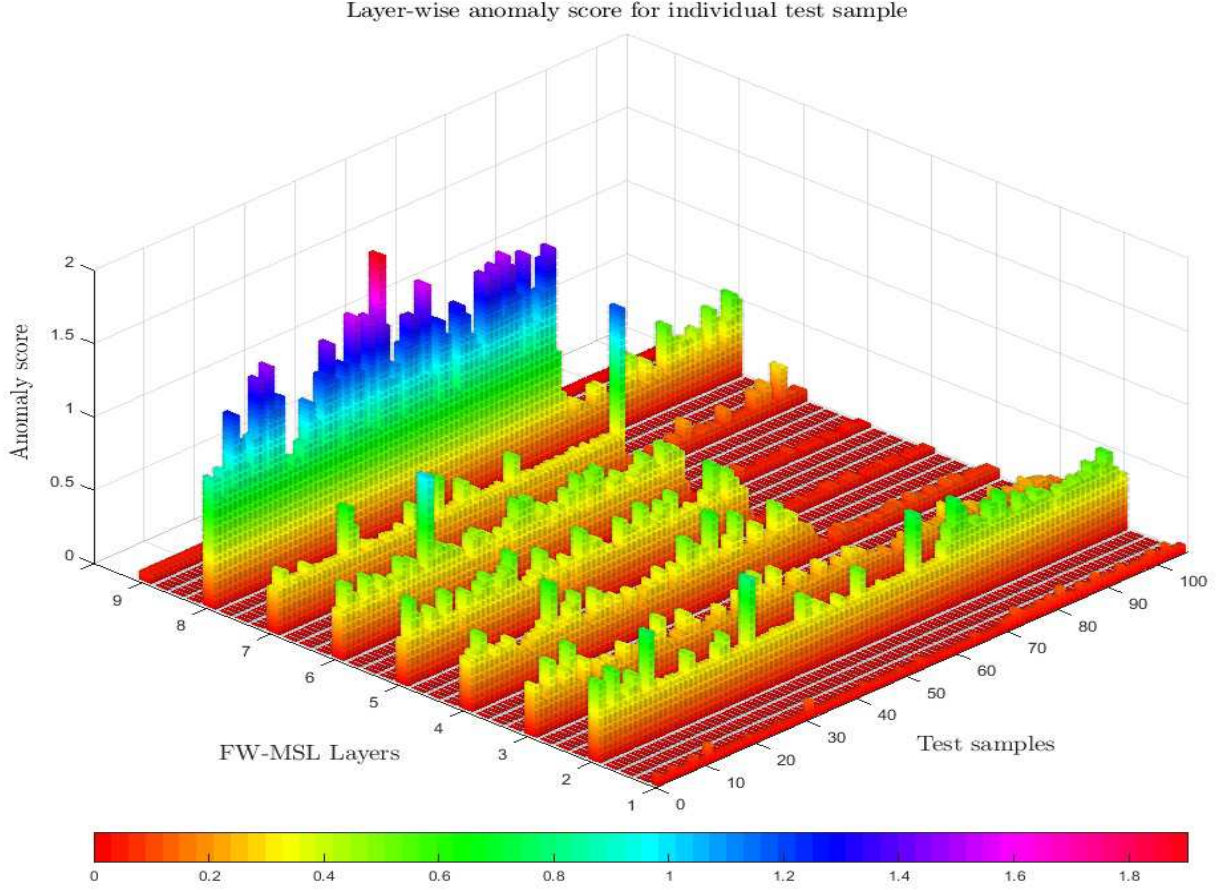


Figure 9. Layer wise reconstruction error computed using 19 is illustrated here. These results also label the species of the trees using a discriminative optimisation routine 12. Spectral and structural anomalies are added in the form of man-made objects and bare ground. Most of these objects are closer to the ground, Layer 7 and 8 clearly show the existence of these anomalies at 15, 20, 34, 69, 103, etc. Samples 34 and 69 record high reconstruction error as this is bare forest floor with some undergrowth. Layer 1 and 9 record the lowest reconstruction error as these layers record mostly background and system noise.

Further analysis is carried out only on the detected anomalous waveforms in order to produce a dense 3D point cloud. Figure 10 illustrates some of these detections. Also, shown is an anomalous waveform and one of the learnt models which is constructed using the learnt dictionaries for each layer. The anomalies, both in the point cloud and waveform have been highlighted with a box. The conifer in Figure 10(A) conceals a large part of a vehicle that has a two surfaces at varying depths. This is obvious in 10(D), highlighted with a box between layer 7 and 8. Subtle anomalies are also detected, e.g., in Figure 10(C) where only a tiny corner of the vehicle falls within the sensor field of view.

5. CONCLUSION & FUTURE WORK

This paper suggests a two phase approach to anomaly detection and target detection in dense cluttered environments. This work is the first piece of work, that the authors are aware of which aims to look for spectral and depth anomalies in raw FW-MSL waveforms. 8 different spectra are introduced in a cluttered forest scenes in the form of man-made objects and vegetation. The suggested approach is widely applicable no matter the sensor modality. This has been considered as a precursor to find *interesting* signals, worthy of more detailed sampling and analysis in order to detect man-made objects, e.g. vehicles under foliage, or mines underwater.

The proposed approach is not restrictive and separate passive hype-spectral imagery registered with Ladar

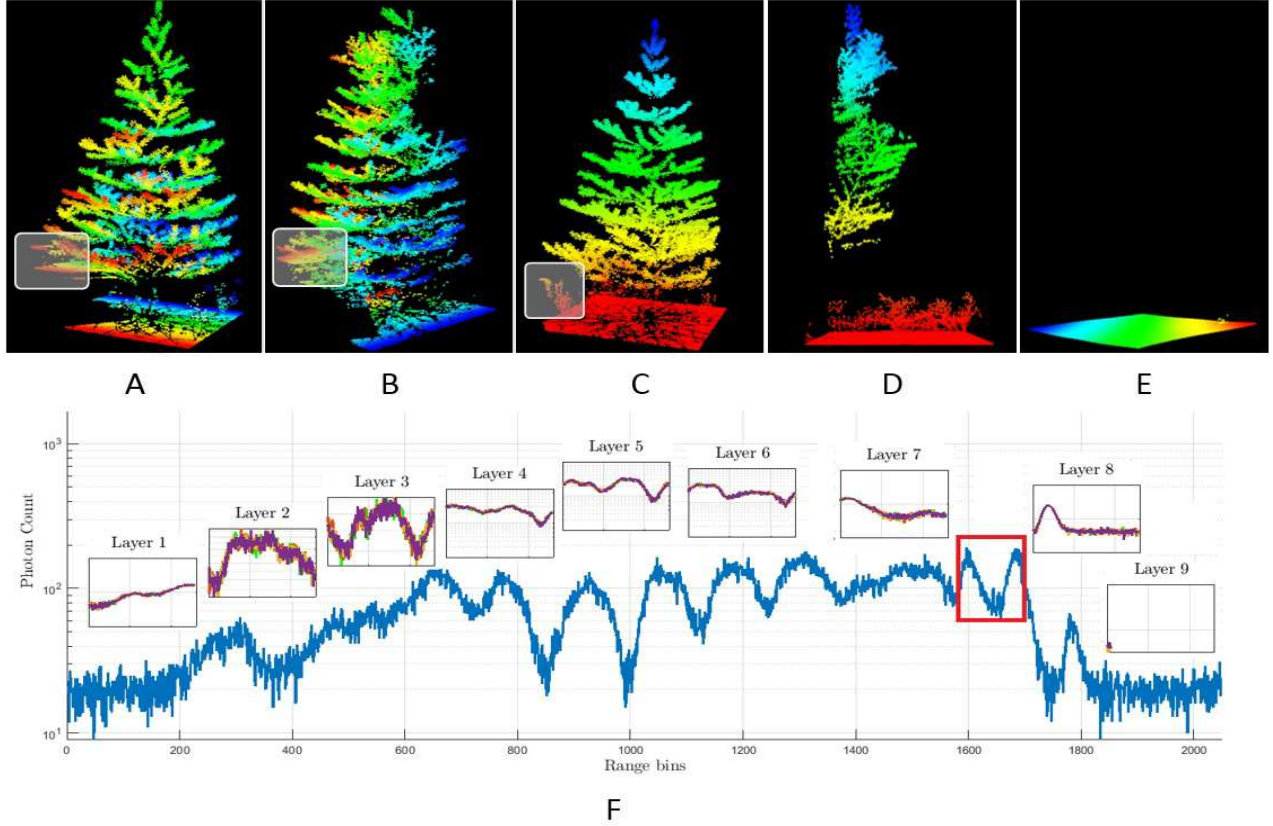


Figure 10. (A) A 3D point cloud of test sample 37, which is detected as an anomaly using the proposed algorithm confirms the presence of large two surfaced object. The detection peaks are shown in between layer 7 and 8. (B) Large section of a man-made object detected on test sample 41. (C) A small subtle anomaly is being detected on layer 8 of test sample 21. (D) and (E) are two anomalies where the ground spectra along with undergrowth dominates the coefficient space. (F) Layer -wise reconstruction of representative models are shown along with test sample 37. The reconstruction is build using dictionaries learnt on individual layers for 32 different wavelengths.

measurements can also be used as long as the training and test measurements are registered or geo-referenced with respect to the sensor. Any number of tree species can be learnt within this framework and as a by-product, the proposed approach allows segmentation of individual foot-prints for semantic scene labelling. This will be explored in detail in future works.

FW-MSL data generation method proposed in this paper generates synthetic data using natural terrain modeller along with a ray-tracer. Comparing this to past MSL work on studying tree physiology,⁷ the synthetic data is as good as real measurements. Future work will concentrate on testing the proposed algorithm on real FW-MSL data where the authors aim to emulate an aerial surveillance scenario with real vehicles and building material concealed under dense foliage. Additional work looking into robust relationship learning algorithms in order to limit the number of false detections could be highly beneficial.

ACKNOWLEDGMENTS

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Platform Grant EP/J015180/1 and the MOD University Defence Research Collaboration in Signal Processing Grant EP/K014277/1.

REFERENCES

- [1] J. Reitberger, P. Krzystek, and U. Stilla, "Analysis of full waveform lidar data for tree species classification," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(Part 3), pp. 228–233, 2006. 1
- [2] N. Brodu and D. Lague, "3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology," *ISPRS Journal of Photogrammetry and Remote Sensing* **68**, pp. 121–134, 2012. 1
- [3] B. Jutzi and U. Stilla, "Range determination with waveform recording laser systems using a Wiener filter," *Journal of Photogrammetry and Remote Sensing* **61**(2), pp. 95–107, 2006. 1
- [4] C. Mallet, F. Bretar, M. Roux, U. Soergel, and C. Heipke, "Relevance assessment of full-waveform lidar data for urban area classification," *ISPRS Journal of Photogrammetry and Remote Sensing* **66**(6), pp. S71–S84, 2011. 1
- [5] N. Chehata, L. Guo, and C. Mallet, "Contribution of airborne full-waveform lidar and image data for urban scene classification," in *16th IEEE International Conference on Image Processing (ICIP), 2009*, pp. 1669–1672, IEEE, 2009. 1
- [6] A. M. Wallace, J. Ye, N. J. Krichel, A. McCarthy, R. J. Collins, and G. S. Buller, "Full wave form analysis for long-range 3D imaging laser radar," *EURASIP Journal on Advances in Signal Processing* **2010**, p. 33, 2010. 1
- [7] A. M. Wallace, A. McCarthy, C. J. Nichol, X. Ren, S. Morak, D. Martinez-Ramirez, I. H. Woodhouse, and G. S. Buller, "Design and evaluation of multispectral lidar for the recovery of arboreal parameters," 2014. 2, 11
- [8] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *Geoscience and Remote Sensing, IEEE Transactions on* **32**(4), pp. 779–785, 1994. 3
- [9] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE* **25**(2), pp. 21–30, 2008. 3
- [10] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, K. Müller, Y.-H. Hu, J. Larsen, E. Wilson, and S. Douglas, "Fisher discriminant analysis with kernels.," in *Neural Networks for Signal Processing*, pp. 41–48, 1999. 7
- [11] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. part i: Greedy pursuit," *Signal Processing* **86**(3), pp. 572–588, 2006. 7
- [12] Y. C. Pati, R. Rezaifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, pp. 40–44, IEEE, 1993. 7
- [13] A. Baldridge, S. Hook, C. Grove, and G. Rivera, "The aster spectral library version 2.0," *Remote Sensing of Environment* **113**(4), pp. 711–715, 2009. 11
- [14] B. B. Mandelbrot, "Fractal aspects of the iteration of $z \lambda z (1-z)$ for complex λ and z ," *Annals of the New York Academy of Sciences* **357**(1), pp. 249–259, 1980. 11